

doi:10.11835/j.issn.1000-582X.2020.01.009

基于改进 EEMD 的卷积神经网络滚动轴承故障诊断

何江江^{1a}, 李孝全^{1b}, 赵玉伟^{1b}, 张保山², 丁海斌³

(1. 空军工程大学 a. 研究生院, b. 防空反导学院, 西安 710053; 2. 92095 部队, 浙江台州 318050;
3. 陆军工程大学训练基地, 江苏徐州 21004)

摘要: 集合经验模态分解(EEMD, ensemble empirical mode decomposition)对信号进行分解, 得到的模态函数(IMF, Intrinsic model function)在 2 端点存在严重的发散现象, 如果将分解结果直接应用到故障诊断系统中, 会导致诊断的准确率下降。首先将支持向量机(SVM, support vector machine)和 EEMD 算法结合进行信号分解, 并利用仿真信号进行可靠性分析; 其次对 SVM(support vector machine)-EEMD 分解的分量进行选择后再分解并构建能量向量, 最后和卷积神经网络结合, 构建滚动轴承故障诊断模型并通过实验验证。结果表明, 改进 EEMD 算法可以有效缓解端点发散问题, 构建的故障诊断模型提高了故障诊断精度。

关键词: 集合经验模态分解; 卷积神经网络; 故障诊断

中图分类号: TN911

文献标志码: A

文章编号: 1000-582X(2020)01-082-08

Fault diagnosis of rolling bearing based on improved EEMD and convolutional neural network

HE Jiangjiang^{1a}, LI Xiaoquan^{1b}, ZHAO Yuwei^{1b}, ZHANG Baoshan², DING Haibin³

(1. a. Graduate School of Air Force Engineering University, b. Air and Missile Defense College of Air Force Engineering University, Xi'an 710053 P. R. China; b. Unit 92095, Taizhou 318050, Zhejiang, P. R. China;
3. Training Base of Army Engineering University of PLA, Xuzhou 21004, Jiangsu, P. R. China)

Abstract: EEMD (ensemble empirical mode decomposition) is an analysis method for signal decomposition. However, there is serious divergence in the two endpoints of its modal function (IMF). If the decomposition results are directly applied to the fault diagnosis system, the diagnosis accuracy will decrease. In the paper, support vector machine (SVM) and EEMD algorithm were combined to decompose signal and the reliability analysis was conducted with simulation signal. After selecting the components of SVM-EEMD decomposition, the signal was decomposed further and the energy vector was constructed. Finally, with a combination of SVM-EEMD and convolutional neural network, rolling bearing fault diagnosis model was constructed and verified by experiment. The experimental comparison results show that the improved EEMD algorithm can effectively solve the problem of the endpoints divergence, and the fault diagnosis model constructed improves the fault diagnosis accuracy.

Keywords: ensemble empirical mode decomposition; CNN; fault diagnosis

收稿日期: 2019-05-23

基金项目: 国家自然科学基金资助项目(51405505)

Supported by National Natural Science Foundation of China(51405505).

作者简介: 何江江(1996—), 男, 硕士研究生, 主要从事电力系统及其自动化方向研究, (Tel)17691350376; (E-mail) 2440772267@qq.com。

通讯作者: 李孝全, 男, 副教授, 硕士生导师, (Tel)13363958159; (E-mail)lxq389@126.com。

滚动轴承在机床设备中被大量使用,起着支撑主轴,承受传递载荷的关键作用^[1],保障主轴可靠稳定运转。机床电机主轴高速、高温等恶劣的工作环境以及轴承自身结构和工艺缺陷使得滚动轴承成为高故障率的零件之一。因此,为保证机床使用寿命和工艺质量,研究滚动轴承故障诊断有重要意义。

机床工作状态下,主轴的高速运转以及其他零部件缺陷或故障(例如:主轴弯曲变形、转子偏心、壳体固定不牢等),使得测量的轴承振动信号中含有许多影响诊断效果的信息,需要对故障特征数据进行提取。目前,特征提取方法主要有快速傅里叶变换(FFT, fast fourier transform)、固有时间尺度分解(ITD, intrinsic time-scale decomposition)、小波包变换(WPT, wavelet packet transform)、模态经验分解(EMD, empirical mode decomposition)等。J.W.库利和 T.W.图基提出的 FFT,减少运算量以实现故障诊断的快速性。对故障振动时间序列进行短时傅里叶变换得到时频谱图,作为卷积神经网络的输入进行训练和故障诊断,同时验证了深度学习对故障诊断的可行性^[1]。但是傅里叶变换使用固定窗口,对非平稳的振动信号不能够得到较高分辨率的频谱图。ITD 是自适应时域分析方法,采用 ITD 将振动时序分解为多个固有旋转分量和单调趋势分量进行故障识别^[2],将 PR 分量进行频谱分析进行故障的预测,但得到的分量波形失真比较严重。文献[3]将振动信号进行经验模态分解,选取具有高能量的前几阶固有模态函数训练 HMM 模型并对轴承故障进行识别,但分解后的信号在两端出现发散现象。

笔者利用支持向量机(SVM)对信号 2 端点进行延伸,再进行 EEMD 分解。然后将分解得到的模式分量作为卷积神经网络的输入,实现滚动轴承的故障诊断。

1 改进 EEMD

EEMD 是 Huang 提出的噪声辅助分解的信号处理算法^[4]。在传统 EMD 的基础上对原始信号加入白噪声,用来弥补传统 EMD 分解存在模态混淆的不足,取得良好的效果。使用 EEMD 分解得到的各个固有模态分量 IMF 在 2 端点仍出现严重的发散现象^[5]。

原始信号 2 端点处没有极值点,导致在进行 3 次样条插值时 2 端点出现很大偏差。对信号分解前使用支持向量机对信号 2 端点进行预测延伸(增加极值点)以缓解端点发散问题。SVM-EEMD 具体步骤如下:

Step1 对原始信号 $X(t)$ 进行 SVM 端点延伸得到信号 $X'(t)$, 增加信号长度。

Step2 在信号 $X'(t)$ 中加入正态白噪声 $N(t)$ 得到信号

$$Y(t) = X'(t) + N(t), \quad (1)$$

Step3 对信号 $Y(t)$ 进行 EMD 分解得到 IMF 分量

$$Y(t) = \sum_{i=1}^m IMF_i(t) + R(t), \quad (2)$$

$R(t)$ 为残余量,代表信号的走势。

Step4 重复前 2 个步骤 m 次,每次加入不同的白噪声,分解得到新分量 $IMF_{ij}(t)$ 。

Step5 由于白噪声具有零均值特性,在平均处理后相互抵消^[6],将 m 次分解得到的 IMF 进行集成平均处理^[7],得到分解结果。

$$IMF_i(t) = \frac{1}{m} \sum_{j=1}^m IMF_{ij}(t), \quad (3)$$

模态分解要保持信号完备性^[8-9],使得分解后的分量相加获得原始信号。例如对信号 $x(t) = \sin(60\pi t) + \sin(15\pi t) \cdot \cos(30\pi t)$ 进行分解和还原。图 1 所示原始信号和分解、还原结果的时域图(蓝色线为分解前,红色线是分解后)。

图 1(a)和 1(c)分别是仿真信号的组成信号和未加 SVM 延伸分解后分量的时域图;图 1(b)和 1(d)是 SVM 延伸分解后分量的时域图,可以看出利用 SVM 对信号进行预测延伸可以很好地缓解端点发散现象;图 1(f)是将分解后分量相加得到的信号时域图,与原信号波形重合。

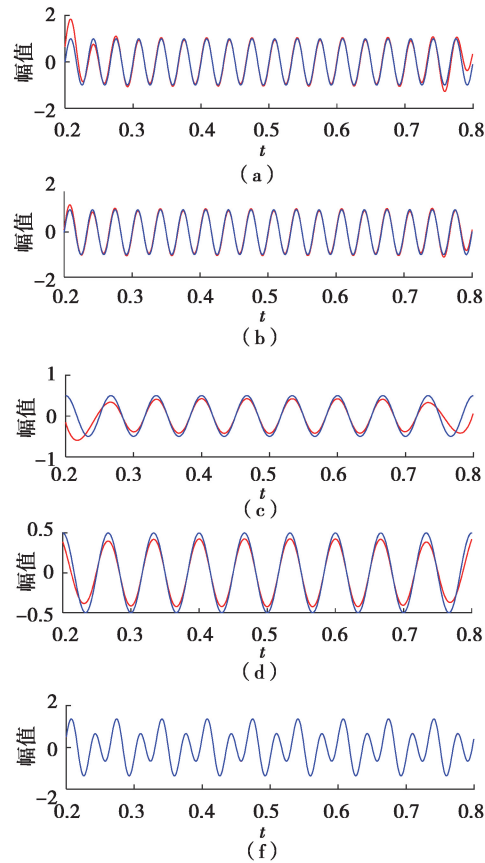


图 1 分解和还原结果

Fig. 1 The results of decomposition and reduction

2 卷积神经网络

卷积神经网络(CNN,convolutional neural networks)是受到生物视觉启发,能够自我提取样本特征的深度神经网络模型,具有平移不变性,在图像处理、人脸识别、特征提取等领域广泛应用^[10]。图 2 所示 CNN 模型包括输入层、卷积层、池化层、全连接层和输出层。

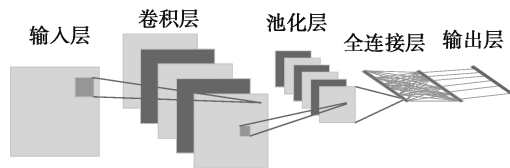


图 2 卷积神经网络组成

Fig. 2 Components of convolutional neural networks

卷积层通过一定大小的卷积核从输入样本中获取特征信息,具有权值共享特性。卷积核(“感受野”)类似于视觉细胞的感受野扫描样本局部区域进行卷积运算^[11]

$$Z(i, j) = \sum_{t=1}^T \sum_{m=1}^{k_c} \sum_{n=1}^{k_c} [X(s_0 i - m, s_0 j - n)W(m, n)],$$

$$i, j = 0, 1, \dots, \frac{H + 2p - k_c}{s_0}, \quad (4)$$

卷积核的 4 个超参数:核数量 T 、核大小 $k_c \times k_c$ 、步长 s_0 、填充大小 p 。卷积核大小和步长的设置可以实现降维,但过大容易造成样本特征缺失,过小导致计算繁琐。将卷积后的值输入到线性整流单元(ReLU

函数)里再加上偏置 b ,提高网络的非线性分割能力。

$$Z_c(i, j) = f(Z(i, j)) + b, \tag{5}$$

池化层对卷积层的输出进行特征筛选,用池化区域中单个点代替整个区域的特征,去除非重要的特征参数。主要有均值池化(求取均值)和最大值池化(求取最大值)2种方法,均是降低特征尺寸获取关键信息。

$$Z_p(x, y) = \max_{1 \leq i, j \leq k_p} (Z_c(i, j)). \tag{6}$$

用单个池化区域最大值代替整个区域特征,类似卷积层,需要适当选取池化窗口大小和步长。

卷积层和池化层是对样本特征的提取过程,全连接层是对特征加权处理过程^[12]。全连接层将池化后的特征转化为向量形式,通过输出函数计算对特征进行分类处理。

卷积神经网络算法包括训练和测试 2 个过程。训练过程包含 2 个阶段:数据前向传播阶段和误差反向传播阶段。前向传播得到正向输出值,通过计算输出值与理想值的误差大小。误差大于期望误差值,根据误差大小对权值进行更新,返回第二步;如果误差小于或等于期望误差值,结束训练。

3 构建故障诊断模型

EEMD 分解方法能实现滚动轴承初始振动数据的去噪处理以及特征分解^[13],而人工神经网络能实现故障分类和智能诊断。所以将 SVM-EEMD 分解与卷积神经网络结合起来实现滚动轴承的故障诊断是高效、准确的,方法的实现过程如图 3 所示。

具体流程如下:

- Step1 故障数据进行分段处理并划分;
- Step2 利用 SVM-EEMD 对处理后数据进行分解;
- Step3 对模态分量 IMFs 进行频谱分析;
- Step4,选择含有大量故障信息的分量作为待分解信号。
- Step5 对选择的分量再进行 SVM-EEMD 分解,得到 n 个分量

并将每个分量分为 m 段;

- Step6 计算 Step5 各分量分段后能量^[14]并归一化;

$$E_{ij} = \sum_{t_1=1}^{t_2} |C_{ij}(t)|^2, \tag{7}$$

$C_{ij}(t)$ 是第 i 个 IMF 分量的第 j 段, t_1 和 t_2 是分段后的时间端点。

- Step7 构造能量特征向量 \mathbf{E} ^[14]

$$\mathbf{E} = [E_{11}, E_{12}, E_{13}, \dots, E_{nm}]. \tag{8}$$

- Step8 训练卷积神经网络并诊断故障类型。

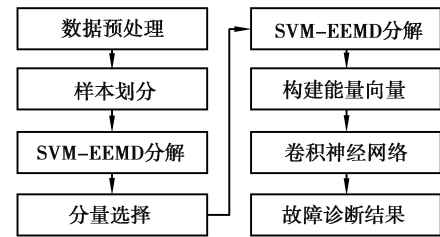


图 3 故障诊断流程

Fig. 3 The flow of Fault Diagnosis

4 实验

4.1 数据来源

实验数据采用美国西储大学轴承数据中心部分数据。利用电火花加工分别对驱动端滚动轴承(型号为 6205-2RS JEM SKF)的内圈、外圈以及滚动体分别设置 7 mils、14 mils、21 mils 单点缺陷故障,其中外圈分别在 3 点钟、6 点钟和 12 点钟设置缺陷。取电动机转速为 1 750 r/min,振动信号采样频率为 12 kHz 下的轴承正常样本和故障样本数据,验证诊断模型可行性。表 1 是驱动端滚动轴承不同位置故障频率(Hz)。

表 1 滚动轴承(6205)故障频率
Table 1 Failure frequency of rolling bearing (6205)

故障位置	内圈	外圈	滚动体
故障频率/Hz	157.94	104.57	68.75

将样本数据集进行分段处理并划分为训练数据集和测试数据集。其中训练数据集共有 210 组,测试数据集共有 70 组。样本的划分和故障的类型如表 2 所示。

表 2 故障类型和样本数据信息

Table 2 The fault type and sample data information

故障类型	训练样本	测试样本	故障类型
正常	6 000×15	6 000×5	0
内圈	6 000×15	6 000×5	1
滚动体	6 000×15	6 000×5	2
外圈 3 点钟	6 000×15	6 000×5	3
外圈 6 点钟	6 000×15	6 000×5	4
外圈 12 点钟	6 000×15	6 000×5	5

4.2 分量选择

对内圈、外圈、滚动体包含的 14 种故障数据分别进行 SVM-EEMD 分解得到幅值较大的 3 个 IMF 分量,再进行频谱解调分析和包络谱分析,选择包含故障特征的分量作为待分解数据。图 4(a)为轴承 7 mils 内圈单点故障原始数据分析结果。图 4(b)是各分量分析结果。时域波形图的横坐标是时间,纵坐标为幅值;频谱波形和包络谱线图的横坐标为频率(Hz),纵坐标为幅值。

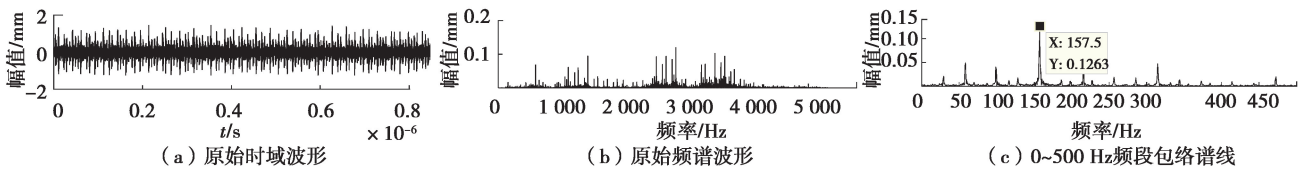


图 4 原始数据分析结果

Fig. 4 The analysis results of original data

从原始数据频谱波形中可以看出数据携带着大量频率在 2 500 Hz~4 000 Hz 的噪声数据,绘制 0~500 Hz 频段的频谱包络曲线,发现数据在 157.5 Hz 出峰值较大,体现出内圈 7 mils 单点故障特征。

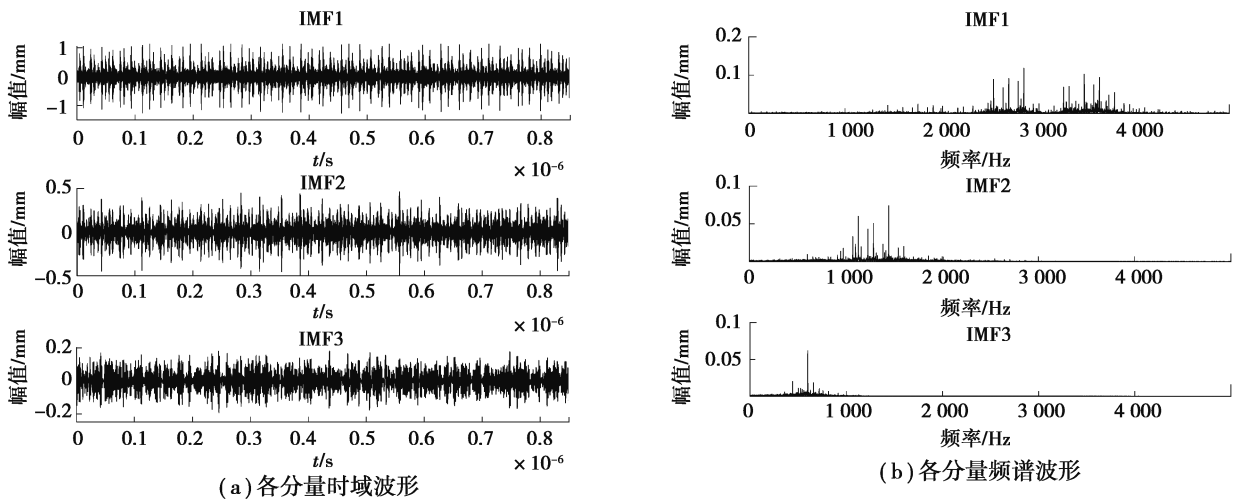


图 5 分量分析结果

Fig. 5 CNNThe results of componential analysis

图 5(a)是轴承 7 mils 内圈单点故障数据分解结果。通过对图 5(b)频谱分析可知分量 IMF3 的频率范围是 0~1 000 Hz,包含着内圈故障的大量信息。所以选择分量 IMF3 再进行 SVM-EEMD 分解,得到 6 个模态函数,将每个分量分为 5 段,并构建能量特征向量(30×1)作为卷积神经网络输入进行训练和测试。

4.3 CNN 诊断结果

将卷积神经网络应用到一维数据处理,将能量特征向量 T 作为网络输入。卷积神经网络共分为 6 层(图 6):输入层(1×30)、卷积层(卷积核大小 1×3×20)、池化层(最大值池化,激活函数为 ReLu,池化后大小为 1×14×20)、卷积层(卷积核大小 1×14×100)、全连接层(100×1,激活函数为 sigmoid)、输出层(6×1, Softmax 函数)。

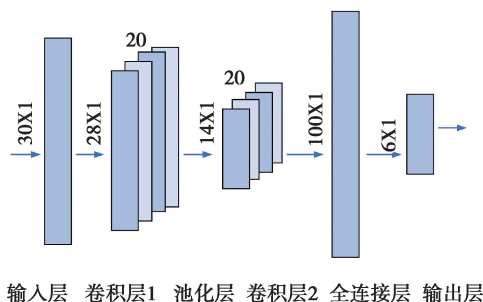


图 6 一维卷积结构

Fig. 6 One-dimensional convolution structure

图 7 所示卷积神经网络训练误差曲线(迭代 50 次)和测试结果的误差值

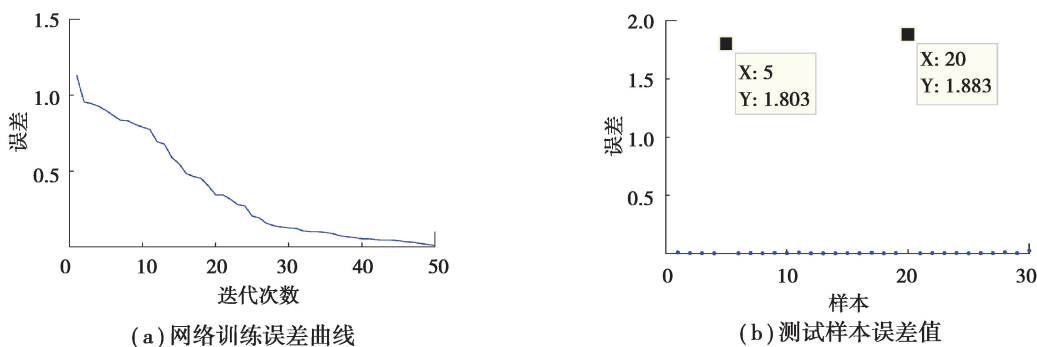


图 7 网络训练和测试误差

Fig. 7 Network training and testing errors

图 7(a)说明训练过程中随着迭代次数增加,误差越来越小。图 7(b)测试结果误差值基本分布在零点附近,而样本 5 和样本 20 的误差值为 1.803 和 1.883,说明测试结果与真实结果相差较远。

表 3 所示 30 个测试样本的测试结果与概率值(输出层使用 softmax 函数)。从表中可以看出第 5、20 个样本真实结果分别为 4、1,而测试结果为 1、4,诊断结果不正确。

表 3 样本测试结果

Table 3 Sample test results

测试类别	真实类别	概率值	测试类别	真实类别	概率值
0	0	0.9 208	3	0	0.9 831
1	1	0.95 524	4	0	0.9 352
2	2	0.9 702	5	0	0.9 839
3	3	0.9 763	0	0	0.9 712

续表3

测试类别	真实类别	概率值	测试类别	真实类别	概率值
1	4	0.91 894	1	0.9 400	
5	5	0.9 803	2	0.9 780	
0	0	0.9 587	3	0.9 653	
1	1	0.98 144	4	0.9 662	
2	2	0.9 401	5	5	0.9 551
3	3	0.9 709	5	0	0.9 873
4	4	0.9 289	1	0.9 709	
5	5	0.9 857	2	2	0.9 639
0	0	0.9 825	3	0.9 211	
1	1	0.9 803	4	4	0.9 620
2	2	0.9 694	2	5	0.8 852

5 结 论

首先将支持向量机与集合模态分解算法结合在一起,对原始数据进行分解。利用仿真信号进行分解验证,可以有效缓解各模态分量在信号端点处的发散现象。然后将 SVM-EEMD 与卷积神经网络结合,提出实现滚动轴承故障诊断模型,通过实验验证可行性和可靠性。诊断结果显示,测试的 30 个样本中有 2 个样本结果与真实结果不符。所以故障诊断模型的诊断精确度(93.33%)比 IEMD+SVM 方法(90.5%)有所提升。

参考文献:

- [1] 李恒, 张氢, 秦仙蓉, 等. 基于短时傅里叶变换和卷积神经网络的轴承故障诊断方法[J]. 振动与冲击, 2018, 37(19): 124-131.
LI Heng, ZHANG Qing, QIN Xianrong, et al. Fault diagnosis method for rolling bearings based on short-time fourier transform and convolution neural network[J]. Journal of Vibration and Shock, 2018, 37(19): 124-131.(in Chinese)
- [2] 权振亚, 张锟, 史晓霞. 基于 ITD 和短时傅里叶变换的故障诊断方法[J]. 装备制造技术, 2018(12): 221-224.
QUAN Zhenya, ZHANG Qian, SHI Xiaoxia. Fault dDiagnosis method based on ITD and short-time fourier transform [J]. Equipment Manufacturing Technology, 2018(12): 221-224.(in Chinese)
- [3] 赵光权, 姜泽东, 胡聪, 等. 基于小波包能量熵和 DBN 的轴承故障诊断[J]. 电子测量与仪器学报, 2019, 33(2): 32-38.
ZHAO Guangquan, JIANG Zedong, HU Cong, et al. Bearing fault diagnosis based on wavelet packet energy entropy and DBN[J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(2): 32-38.(in Chinese)
- [4] 陈雪娇, 仇满意, 赵文涛. 基于 EEMD 信号处理的滚动轴承故障诊断[J]. 技术与市场, 2019, 26(3): 121, 123.
CHEN Xuejiao, QIU Manyi, ZHAO Wentao. Rolling bearing fault diagnosis based on EEMD signal processing [J]. Technology & Market, 2019, 26(3): 121, 123.(in Chinese)
- [5] Jiali Z, Yuan L, Junwei L, et al. Research of mechanical fault SVM intelligent recognition based on EEMD sample entropy [J]. Sensors & Transducers, 2014, 179(9):141-148.
- [6] Weili Q, Wenjin Z, Chen L. Rolling bearing fault diagnosis based on ensemble empirical mode decomposition, information entropy and random forests[J]. Vibroengineering Procedia, 2015, 5:211-216.
- [7] Long H, Chengwei L, Hongchen L. Feature extraction method of rolling bearing fault signal based on EEMD and cloud model characteristic wntropy[J]. Entropy, 2015, 17(10): 6683-6697.

- [8] Zheng K, Luo J F, Zhang Y, et al. Incipient fault detection of rolling bearing using maximum autocorrelation impulse harmonic to noise deconvolution and parameter optimized fast EEMD[J]. ISA Transactions,2019,89:256-261.
- [9] Cheng Y, Wang Z W, Chen B Y, et al.An improved complementary ensemble empirical mode decomposition with adaptive noise and its application to rolling element bearing fault diagnosis[J]. ISA transactions, 2019,91:218-234.
- [10] Shisheng Z, Song F, Lin L. A novel gas turbine fault diagnosis method based on transfer learning with CNN[J]. Measurement,2019,137:435-453.
- [11] Shu L H, Ng Eddie Y K, Tan R S, et al. Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats[J]. Computers in biology and medicine,2018,102:278-287.
- [12] Han T, Liu C, Wu L J, et al. An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems [J]. Mechanical Systems and Signal Processing,2019,117:170-187.
- [13] Jiang H, Cai Q, Zhao H, et al. Rolling bearing fault feature extraction under variable conditions using hybrid order tracking and EEMD[J]. Journal of Vibroengineering, 2016, 18(7):4449-4457.
- [14] Wu D, Huimin Z, Xinhua Y, et al. A fault feature extraction method for motor bearing and transmission analysis [J]. Symmetry, 2017, 9(5):60.
- [15] 徐可, 陈宗海, 张陈斌, 等. 基于经验模态分解和支持向量机的滚动轴承故障诊断[J]. 控制理论与应用, 2019(6): 915-922.
- XU Ke, CHEN Zonghai, ZHANG Chenbin, et al. Fault diagnosis of rolling bearing based on empirical mode decomposition and support vector machine[J]. Control Theory & Applications, 2019(6):915-922.(in Chinese)

(编辑 侯 湘)